



TOOLKIT

COM PUC AVALUAR L'IMPACTE D'UN PROGRAMA O INTERVENCIÓ

MARÇ 2024

IMPULSEN:

ivàlua  Institut Català d'Avaluació
de Polítiques Públiques



Taula d'entitats
del Tercer Sector Social
de Catalunya

ÍNDEX

Què és?	3
Per a què serveix?	4
Com aplicar l'eina?	6
Com avaluar l'impacte d'un programa?	6
Experiment aleatori controlat (o mètode experimental).....	8
Els mètodes quasiexperimentals	9
<i>El mètode del matching o emparellament</i>	10
<i>El mètode de diferència en diferències (o doble diferència)</i>	11
<i>El disseny de regressió discontinua</i>	13
Què cal tenir en compte?	14
Exemples	16
Per saber-ne més	17

QUÈ ÉS?

Un programa, intervenció o política pública té com a objectiu principal mitigar o solucionar una problemàtica o situació social insatisfactòria. Per exemple, un programa que proporciona cursos de formació en habilitats professionals a persones joves en situació d'atur té com a objectiu principal incrementar la inserció laboral de les persones participants. Les baixes taxes d'ocupació entre la població jove són la problemàtica social que dona raó de ser al programa. La reducció d'aquestes taxes, a través de la inserció laboral de les participants, és l'objectiu principal del programa de formació.

Una avaluació d'impacte d'un programa o intervenció, mitjançant una anàlisi estadística i emprant dades quantitatives, busca determinar si el programa ha estat efectiu en la consecució del seu objectiu principal. És a dir, si el programa ha assolit mitigar o solucionar la problemàtica social insatisfactòria que li dona raó de ser. En el nostre exemple, una avaluació d'impacte del programa de formació busca determinar si el programa ha aconseguit incrementar la inserció laboral de les participants.

Un element fonamental d'una avaluació d'impacte és el concepte de causalitat. **Una avaluació d'impacte busca determinar si els canvis en la situació social insatisfactòria que el programa vol mitigar es poden atribuir de forma única al programa.** Dit d'una altra manera, una avaluació d'impacte busca donar resposta a la pregunta: quin ha estat l'efecte causal del programa en la mitigació de la problemàtica social que aquest busca solucionar? Continuant amb el mateix exemple: quin ha estat l'efecte causal del programa de formació en la inserció laboral de les persones participants? Donat que els fenòmens socials es veuen afectats per múltiples factors, determinar si els canvis en aquests fenòmens són la conseqüència única (l'efecte causal) del programa objecte d'avaluació és el repte metodològic més important en una avaluació d'impacte.

Una eina molt útil per entendre i delimitar els objectius d'una avaluació d'impacte és la Teoria del Canvi (TdC) del programa (vegeu Eina 1 del Toolkit). La TdC posa en relació, de forma clara i explícita, els diferents elements que conformen el programa o intervenció: descriu la problemàtica social que dona raó de ser al programa, els recursos dels quals es disposa per a implementar-lo, les activitats que aquest porta a terme, els productes (o *outputs*) que es deriven de forma directa d'aquestes activitats, i els resultats (o *outcomes*) que el programa pretén afectar en última instància. D'aquesta forma, la TdC exposa de quina forma el programa serà capaç de solucionar la problemàtica social que el motiva.

En aquest sentit, és important diferenciar els productes (*outputs*) dels resultats (*outcomes*). Els *outputs* d'un programa són aquells productes o serveis que resulten de forma directa de l'execució de les activitats del programa (quan aquestes es despleguen correctament). Els *outcomes* són els resultats finals que el programa vol afectar en última instància, i que mesuren aspectes de la problemàtica social que es vol solucionar. **Una avaluació d'impacte busca determinar si el programa objecte d'avaluació ha tingut un efecte causal en l'outcome o outcomes.** D'aquesta manera, estarem determinant si el programa aconsegueix resoldre la problemàtica social que li dona raó de ser.

En el programa de formació per a joves en situació d'atur, un *output* del programa seria, per exemple, la participació dels i les joves en la formació. Els *outcomes* del programa podrien ser el fet de trobar feina, després de la realització dels cursos, o el salari que obtenen en aquesta feina. D'acord amb l'explicació anterior, una avaluació d'impacte busca determinar si el programa de formació ha tingut un efecte causal en la probabilitat dels joves participants de trobar una feina o en els canvis en el seu salari.

PER A QUÈ SERVEIX?

Com s'ha comentat, la pregunta principal a la qual una avaluació d'impacte vol donar resposta és: el programa ha assolit solucionar la problemàtica social que li dona raó de ser? La resposta a aquesta pregunta és fonamental amb relació al moviment conegut com a *política basada en l'evidència*. Aquest moviment, que està guanyant cada cop més presència arreu del món, proposa que les decisions a l'hora d'aplicar intervencions socials estiguin basades en informació provinent d'anàlisis rigoroses sobre la mesura en què aconsegueixen els seus objectius. En aquest sentit, l'avaluació d'impacte és fonamental per tal de proporcionar informació sobre quins programes funcionen i quins no, i contribuir d'aquesta manera a informar les decisions sobre l'assignació dels recursos de les organitzacions que es dediquen a dissenyar i implementar programes socials.

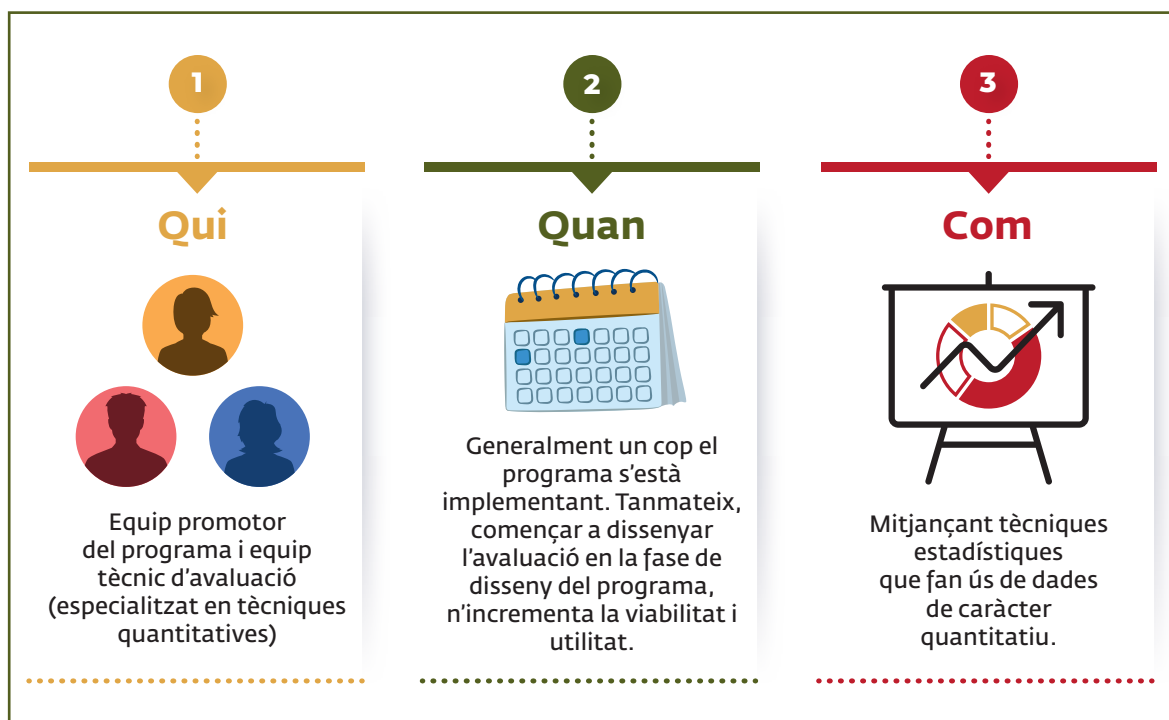
Les principals decisions que una avaluació d'impacte pot ajudar a prendre són:

- › **Determinar si s'ha d'implementar un nou programa.** Les problemàtiques socials són canviants. És habitual, per tant, que programes que s'implementen per solucionar una determinada problemàtica ja no resultin efectius perquè el context social ha canviat. També és usual que sorgeixin noves

problemàtiques per a les quals no existeix evidència sobre quins programes o intervencions són efectius. En aquest cas, una avaluació d'impacte que analitzi l'efectivitat d'un nou programa proporciona informació molt útil per decidir si donar continuïtat al nou programa o escalar la seva implementació a més territoris o a un conjunt més ampli de població. Quan s'implementen nous programes, és habitual començar implementant un programa pilot, el qual s'aplica només a territoris concrets o a un subconjunt de la població. En aquests casos, per tant, és molt recomanable acompanyar la implementació del programa d'una avaluació d'impacte.

- › **Determinar si donar continuïtat a un programa establert.** Existeixen programes que fa temps que s'implementen i que, sigui perquè el context social ha canviat o perquè no s'ha basat la seva implementació en evidència, no són efectius per a la solució del problema. En aquests casos, una avaluació d'impacte, si determina que el programa no és efectiu, pot servir per prendre la decisió de deixar de destinar recursos a la seva implementació, destinant-los a altres programes que hagin demostrat major efectivitat.
- › **Determinar quin tipus d'implementació és més efectiva dins d'un mateix programa.** Dins d'un mateix programa poden existir diverses modalitats o línies d'implementació. Per exemple, en un programa que proporciona aliments a famílies en risc d'exclusió social, es pot fer la provisió d'aquests aliments mitjançant l'entrega directa d'una bossa d'aliments o mitjançant una targeta de prepagament. En aquests casos, una avaluació d'impacte permet determinar quina modalitat d'implementació és més efectiva en la consecució de l'objectiu final del programa. Aquesta informació serveix de base per a modificar el tipus d'implementació i incrementar d'aquesta manera l'efectivitat del programa.

COM APLICAR L'EINA?



COM AVALUAR L'IMPACTE D'UN PROGRAMA? ELS MÈTODES D'AVAUACIÓ D'IMPACTE

Com s'ha comentat, la causalitat és l'element fonamental de les avaluacions d'impacte. Una avaluació d'impacte busca determinar quins canvis s'han produït en els *outcomes* d'interès a *conseqüència* de la implementació del programa objecte de l'avaluació. O, dit d'una altra manera, busca respondre a la pregunta: Quin és l'efecte causal del programa en l'*outcome* d'interès? Com veurem, poder determinar si els canvis en l'*outcome* d'interès són conseqüència (són els *efectes causals*) del programa és, a l'hora, el principal atractiu i el principal repte metodològic de les avaluacions d'impacte.

Per a poder determinar quin és l'*efecte causal* d'un programa o intervenció en un *outcome* d'interès, una avaluació d'impacte utilitza el que s'anomena "mètodes d'avaluació d'impacte", que són tècniques estadístiques que utilitzen dades de caràcter quantitatiu. L'atractiu principal d'aquestes tècniques és que, **si es compleixen diversos supòsits i es disposa de suficient informació estadística, són capaces de determinar l'efecte causal del programa**, és a dir,

aquells canvis que s'han produït en l'*outcome* d'interès *a conseqüència* de la implementació del programa.

Per a entendre com els mètodes d'avaluació d'impacte són capaços de determinar l'*efecte causal* d'un programa, és molt útil entendre el concepte de *contrafactual*. Imaginem que volem conèixer quin és l'efecte causal del programa de formació en la probabilitat dels joves participants de tenir una feina. Imaginem, també, que disposem d'una mostra de joves que han participat en el programa de formació, i que podem observar el percentatge de joves que tenen feina després de participar en el programa, que és, per exemple, d'un 80%. Quin és l'efecte causal de participar en el programa en la probabilitat de tenir una feina? Per a poder determinar aquest efecte, els mètodes d'avaluació d'impacte busquen respondre a la pregunta: què hauria passat amb el jovent que ha participat en el programa si aquest mateix jovent *no hi hagués participat*? Si poguéssim saber el percentatge de joves que tindrien feina si aquests no haguessin participat en el programa (posem, un 70%), l'efecte causal de participar en el programa en la probabilitat de tenir una feina seria de $80-70=10\%$. És a dir, podríem dir que participar en el programa incrementa la probabilitat de tenir una feina en 10 punts percentuals. Evidentment, no podem veure el percentatge de joves que tenen feina si aquests no haguessin participat en el programa perquè aquests joves en realitat sí que hi han participat. Aquesta mesura es coneix com a *contrafactual*. És a dir, el *contrafactual* és la mesura de l'*outcome* d'interès per al grup de persones que participen en un programa si aquestes persones no hi haguessin participat.

Com veurem a continuació, els mètodes d'avaluació d'impacte utilitzen tècniques estadístiques per a poder calcular (o estimar) el *contrafactual*. Si podem calcular aquest *contrafactual* de forma robusta, l'impacte d'un programa és simplement la diferència entre allò que ha passat amb les persones participants en el programa (que podem observar de forma directa) i allò que hauria passat si no hi haguessin participat (el *contrafactual*). Imaginem que, a part del percentatge de joves amb feina del grup de joves que ha participat en el programa, disposem també d'informació sobre el percentatge de joves amb feina d'un grup de joves que no ha participat al programa. En termes tècnics, diem que tenim un grup de tractament (persones que han participat al programa) i un grup de comparació (persones que no han participat en el programa). Si podem demostrar que el grup de tractament és idèntic al grup de comparació respecte a les seves característiques (edat, sexe, nivell educatiu, etc.), de forma que l'única diferència entre ambdós grups és que el primer ha participat en el programa i el segon no, una comparació entre l'*outcome* d'interès dels dos grups (el per-

centatge de joves que tenen feina) ens donarà una estimació de l'efecte causal del programa. **Trobar un grup de comparació el més semblant possible al grup de tractament és, precisament, el que busquen els diferents mètodes d'avaluació d'impacte.** Com millor se'n surti el mètode en aquesta tasca, millor serà en estimar l'efecte causal del programa de forma robusta.

A continuació, s'exposen els diferents mètodes que es poden utilitzar per a estimar l'impacte d'un programa o intervenció en un *outcome* d'interès. Com s'ha comentat, l'objectiu principal és trobar un grup de comparació el més semblant possible al grup de tractament per comparar l'*outcome* d'interès entre els dos grups.

EXPERIMENT ALEATORI CONTROLAT (O MÈTODE EXPERIMENTAL)

L'experiment aleatori controlat (en anglès, *randomized control trial* o RCT) és considerat el *gold standard* dels mètodes d'avaluació d'impacte. Amb això ens referim al fet que és el mètode que aconsegueix estimar l'efecte causal d'un programa de forma més creïble o robusta. La raó rau en el fet que la tècnica que usa per a construir el grup de comparació -l'aleatorització del programa-, assoleix que aquest grup de comparació sigui pràcticament idèntic al grup de tractament en totes les característiques de les persones, excepte en la participació en el programa.

Per a construir un grup de tractament i un grup de comparació, un experiment aleatori determina si un individu participa o no al programa mitjançant una assignació aleatòria. El mètode és molt utilitzat per a avaluar l'efectivitat de productes farmacèutics, especialment medicaments. En aquest cas, per a determinar si un medicament és efectiu a l'hora de mitigar un determinat problema de salut, el medicament es proporciona a un grup de tractament determinat de forma aleatòria. Les persones a les quals, també de forma aleatòria, no se'ls proporciona el medicament, formen el grup de comparació. Com que l'assignació de les persones al grup de tractament i al grup de comparació s'ha fet de forma aleatòria, els dos grups seran idèntics en totes les característiques individuals a excepció que un haurà rebut el medicament i l'altre no. La comparació de l'*outcome* d'interès entre els dos grups, per tant, serà una mesura robusta de l'efecte causal del medicament. Prova de la robustesa d'aquest mètode a l'hora d'estimar efectes causals n'és el fet que les diferents agències nacionals i

internacionals encarregades d'aprovar la comercialització de medicaments en requereixen l'aplicació.

Tot i que l'aplicació del mètode experimental per avaluar l'impacte de programes socials és menys habitual que en l'àmbit sanitari, aquest està guanyant cada cop més presència. Aprofitar el fet que, sovint, per falta de recursos, els programes socials no poden aplicar-se des d'un inici a tota la població potencialment beneficiària, ens permet dissenyar a l'inici del programa un experiment aleatori controlat per a respondre de forma robusta preguntes sobre l'efectivitat del programa. El fet de planejar l'avaluació d'impacte abans d'implementar el programa, permet ser molt concret respecte a quines preguntes volem donar resposta amb l'avaluació, la qual cosa incrementa molt la utilitat d'aquesta per a la presa de decisions.

ELS MÈTODES QUASIEXPERIMENTALS

Com s'ha comentat, l'experiment aleatori controlat és el mètode més robust per avaluar l'impacte d'un programa, donat que aconseguix que els dos grups siguin idèntics en totes les característiques individuals a excepció de la participació en el programa. La seva aplicació, però, només és possible si l'avaluació es dissenya abans d'implementar el programa. Com ho fem, doncs, per avaluar l'impacte d'un programa que ja està en funcionament i pel qual no hi ha hagut una assignació aleatòria de la participació? En aquests casos s'utilitzen els mètodes quasiexperimentals, els quals utilitzen tècniques diferents de l'aleatorització per construir un grup de comparació. Com el seu nom suggereix, cadascun d'aquests mètodes fa servir una tècnica diferent per a construir un grup de comparació que sigui el més semblant possible a aquell grup que hauria sorgit de l'aplicació de l'aleatorització.

A continuació, es descriuen els possibles mètodes quasiexperimentals que es poden aplicar per a avaluar l'impacte d'un programa quan l'aleatorització de la participació no ha estat possible. Com veurem, la possibilitat d'aplicar cada un dels mètodes dependrà de les característiques del programa objecte d'avaluació i de la disponibilitat de dades.

EL MÈTODE DEL *MATCHING* O EMPARELLAMENT

El mètode del *matching* emprà un algoritme per a construir un grup de comparació artificial. Per aplicar el mètode, partim d'una mostra de persones que ha participat en el programa (grup de tractament), i d'una mostra de persones que no hi ha participat (grup de comparació potencial). **Per a cadascuna de les persones del grup de tractament, l'algoritme busca la o les persones del grup de comparació potencial que més s'hi assemblin en funció d'un conjunt de característiques de les persones que podem observar i mesurar** (com l'edat, el sexe, el nivell educatiu, etc.). El conjunt de persones del grup de comparació potencial seleccionades mitjançant aquest algoritme formaran el grup de comparació. Mitjançant aquest mètode, garantim que el grup de comparació sigui igual al grup de tractament en relació amb les característiques que hem inclòs en l'algoritme (l'edat, el sexe, el nivell educatiu, etc.). És a dir, de mitjana, les persones del grup de tractament i del grup de comparació tindran la mateixa edat, el mateix sexe, el mateix nivell educatiu, etc.

Igual que fem en el mètode experimental, la diferència en l'*outcome* d'interès entre el grup de tractament i aquest grup de comparació "artificial" ens donarà una mesura de l'impacte del programa.

A diferència del mètode experimental, però, el mètode del *matching* presenta un problema important. Per a entendre-ho, imaginem que volem estimar l'impacte del programa de formació per a joves aturats en la probabilitat que tinguin una feina mitjançant el mètode del *matching*. Per a fer-ho, partim d'una mostra de joves que han participat en el programa i d'una mostra de joves que no hi ha participat. Per a construir el grup de comparació, apliquem l'algoritme del *matching* utilitzant un conjunt de característiques dels joves que podem veure i mesurar (com l'edat, el sexe, el nivell educatiu i l'historial laboral). D'aquesta manera, les persones del grup de comparació i les del grup de tractament tindran, de mitjana, la mateixa edat, el mateix sexe, el mateix nivell educatiu i el mateix historial laboral.

Imaginem, però, que hi ha una característica dels joves que no podem observar i que, per tant, no podem incloure en l'algoritme, com per exemple el nivell de motivació dels joves a l'hora de buscar feina. En aquest cas, és molt probable que el jovent que ha participat en el programa estigui més motivat a l'hora de buscar feina que el que no hi ha participat. En conseqüència, com que no hem inclòs aquesta variable en el nostre algoritme, el jovent del grup de comparació estarà menys motivat. Per altra banda, és també molt probable que els/les joves que tenen una motivació més alta a l'hora de buscar feina tinguin més èxit

a l'hora de trobar-la. Si aquest és el cas, l'efecte del programa que obtindrem de l'aplicació del mètode del *matching* serà esbiaixat, és a dir, no serà l'efecte causal real del programa. La raó és que l'efecte que trobem no estarà únicament causat per la participació en el programa, sinó també perquè el jovent que ha participat en el programa tenia, abans de l'inici del programa, una motivació més alta a l'hora de buscar feina.

Aquest és el problema principal del mètode del *matching*, perquè a l'hora d'avaluar l'impacte d'un programa, és habitual que no disposem d'informació sobre característiques rellevants de les persones. Si aquestes característiques que no podem mesurar (com la motivació, la intel·ligència o diversos trets de personalitat) influeixen en la probabilitat de participar en el programa i en l'*outcome* d'interès, el mètode del *matching* no serà capaç d'estimar l'efecte causal real del programa. En aquests casos, haurem de recórrer a altres mètodes -descrits a continuació-, que busquen solucionar aquest problema.

EL MÈTODE DE DIFERÈNCIA EN DIFERÈNCIES (O DOBLE DIFERÈNCIA)

Si es compleixen certes condicions que s'explicaran més endavant, el mètode de diferència en diferències ens permetrà solucionar el problema exposat. Per aplicar el mètode de diferència en diferències, partim, igual que en el mètode del *matching*, d'un grup de persones que han participat en el programa (grup de tractament) i d'un grup de persones que no hi ha participat (grup de comparació). Igual que amb el *matching*, podem observar l'*outcome* d'interès en els dos grups després del programa. A diferència del *matching*, però, disposem d'una informació addicional: l'*outcome* d'interès per als dos grups abans de l'inici del programa. A partir d'aquí, per a calcular l'efecte del programa, el mètode de diferència en diferències **calcula, en primer lloc, la diferència de l'*outcome* d'interès entre el grup de tractament i el grup de comparació després del programa**. Aquesta magnitud és, com hem vist, l'estimació de l'efecte del programa que obteníem aplicant el mètode del *matching*. **Per altra banda, calculem la diferència de l'*outcome* d'interès entre el grup de tractament i el grup de comparació abans de l'inici del programa. L'estimació de l'efecte del programa serà, simplement, la diferència entre aquestes dues magnituds** (la diferència de la diferència).

Per a entendre perquè el mètode de diferència en diferències permet resoldre el problema que ens trobàvem amb el mètode del *matching* és útil tornar a l'exem-

ple del programa de formació per a joves en situació d'atur. Recordem que, mitjançant el mètode del *matching*, havíem calculat l'efecte del programa calculant la diferència en el percentatge de joves amb feina després del programa entre el grup de tractament i el grup de comparació. Però, com que les persones del grup de tractament, abans del programa, tenien nivells de motivació més alts en el moment de buscar feina, aquest càlcul no permetia obtenir una estimació real de l'efecte causal del programa. Per a solucionar aquest problema, el mètode de diferència en diferències fa un càlcul addicional. A aquesta primera diferència calculada després del programa li resta la diferència entre els dos grups en el percentatge de joves amb feina observada *abans* del programa. Com que la diferència entre els dos grups en el percentatge de joves amb feina observada abans del programa és causada per la diferència en la motivació a l'hora de buscar feina, aquest càlcul addicional ens permet *controlar* per aquest efecte en la nostra estimació de l'efecte causal del programa, produint d'aquesta manera estimacions no esbiaixades.

l aquest és l'atractiu principal del mètode de diferència en diferències. Si existeixen, com acostuma a ser el cas, característiques de les persones que no podem analitzar i que afecten alhora la probabilitat de participar en el programa i l'*outcome* d'interès, el mètode de diferència en diferències és capaç de *controlar* per aquestes característiques, encara que no les puguem analitzar. Per això diem que el mètode de diferència en diferències és més robust que el mètode del *matching*. En aquests casos, permet solucionar aquest problema i d'aquesta manera és capaç d'estimar l'efecte causal real del programa.

Cal tenir en compte, però, que el mètode de diferència en diferències no sempre serà capaç d'estimar l'efecte causal real del programa. La raó rau en el fet que, com hem vist, aquest mètode és capaç de *controlar* per característiques no observables com la motivació, la intel·ligència o diversos trets de personalitat de les persones. Però el mètode és capaç de controlar aquestes característiques només si aquestes *no canvien en el temps* o, més concretament, si aquestes *no canvien durant el transcurs del programa*. És raonable pensar que característiques de les persones com la motivació o la intel·ligència no canvien durant el temps en el qual s'implementa un programa, però què passa amb altres variables, com poden ser el cicle econòmic o l'existència d'altres programes socials contemporanis al que estem avaluant? És possible, per exemple, que el cicle econòmic afecti de forma diferent al grup de tractament i el grup de comparació. Si, com és també probable, el cicle econòmic afecta la probabilitat que els joves tinguin una feina, el mètode de diferència en diferències no serà capaç de *controlar* per aquesta variable i, per tant, no serà capaç de proporcionar estimacions no es-

biaixades de l'efecte causal del programa. En aquests casos, haurem de recórrer a altres mètodes, com el que s'exposa a continuació.

EL DISSENY DE REGRESSIÓ DISCONTÍNUA

El disseny de regressió discontinua és un mètode d'avaluació d'impacte molt robust, però que només es pot aplicar quan el programa objecte d'avaluació presenta una característica molt concreta: que el mecanisme que es fa servir per a decidir qui participa en el programa estigui basat en l'assignació a partir d'una variable contínua (una característica de les persones o unitats potencialment receptors del programa) i un llindar en aquesta variable. Tot i que molts programes socials no presenten aquesta característica, no és estrany trobar-ne que sí que ho fan. Per exemple, alguns programes destinats a mitigar situacions de pobresa es proporcionen a persones o famílies que estan per sota d'un llindar de renda determinat. Alguns programes destinats a millorar el rendiment educatiu es proporcionen a alumnes o a escoles per sota d'una nota concreta en relació amb resultats en exàmens. Alguns programes, com per exemple polítiques actives d'ocupació o programes de pensions, es proporcionen només a persones que tenen una edat determinada. O programes destinats a empreses es proporcionen a aquelles que estan per sota o per sobre d'un determinat nombre de persones treballadores.

El disseny de regressió discontinua utilitza aquest llindar per a construir el grup de tractament i el grup de comparació. Imaginem que, en l'exemple del programa de formació per a joves en situació d'atur, la formació es proporciona només a joves que tenen un nivell de renda anual inferior a 20000 euros. En aquest cas, el grup de tractament estaria format per persones que estan just per sota d'aquest llindar (per exemple, persones amb un nivell de renda entre 17000 i 20000 euros), i el grup de comparació estaria format per persones que estan just per sobre (de 20001 a 23000 euros). L'atractiu principal d'aquest mètode és que, si es compleixen certes condicions, variacions en la variable que determina l'assignació (en aquest cas, el nivell de renda) que es produeixen just al voltant del llindar no estan relacionades ni amb les altres característiques de les persones ni amb l'*outcome* d'interès. Però aquestes variacions que es produeixen just al voltant del llindar sí que determinen qui participa i qui no participa en el programa. En conseqüència, podem dir que en aquest cas la participació en el programa és com si s'hagués fet de forma aleatòria, com en el mètode experimental. Per tant, la diferència en l'*outcome* d'interès entre el grup de tractament i el grup de comparació ens donarà una estimació no esbiaixada de l'efecte causal del programa.

QUÈ CAL TENIR EN COMPTE?

► Cal haver construït o analitzat una Teoria del Canvi (TdC) abans de dissenyar l'avaluació

Recordem que una TdC del programa descriu com se suposa que aquest assolirà els canvis pels quals s'ha dissenyat. Per tal d'incrementar la utilitat de l'avaluació d'impacte, és important haver analitzat la seva TdC abans de dissenyar l'avaluació. Això ens ajudarà a formular preguntes d'avaluació rellevants, a escollir els *outcomes* per als quals volem analitzar l'efecte i a saber quines bases de dades necessitarem per a avaluar.

► Cal conèixer com ha funcionat la implementació

Si l'avaluació d'impacte determina que el programa no aconsegueix assolir els seus objectius, però aquest no s'ha implementat com estava dissenyat, estarem concloent erròniament que el programa està mal dissenyat. Per altra banda, si l'avaluació d'impacte determina que el programa té efectes positius, saber que el programa s'ha implementat com estava dissenyat reforça les conclusions de l'avaluació d'impacte. Per tant, si el procés d'implementació del programa no s'ha analitzat prèviament, resulta d'utilitat per tal d'interpretar correctament els resultats de l'avaluació d'impacte acompanyar aquesta amb una avaluació de la implementació.

► Cal deixar suficient temps perquè els impactes s'hagin pogut produir

En general, els impactes esperats dels programes socials no són immediats, sinó que tarden un cert temps a produir-se. Cal, per tant, tenir el coneixement de quan s'espera que els impactes s'esdevinguin abans d'iniciar una avaluació d'impacte. Si el coneixement ens indica que encara no ha passat prou temps perquè els impactes esperats s'hagin pogut produir, és recomanable posposar l'avaluació per al moment en què aquests hagin pogut esdevenir-se.

► Dissenyar l'avaluació d'impacte en el moment de disseny del programa maximitza la viabilitat i la utilitat de l'avaluació

L'avaluació d'impacte s'ha de dur a terme quan el programa ja està en funcionament, un cop els impactes s'han pogut produir i s'ha pogut recollir la informació necessària. Començar a dissenyar l'avaluació en el moment en què es dissenya el programa, però, resulta extremadament important per tal d'incrementar les possibilitats que aquesta sigui viable i útil per a la presa de decisions. En primer lloc, fer una avaluació d'impacte requereix l'aplicació d'un mètode d'avaluació. Com s'ha vist, perquè el mètode d'avaluació sigui viable i robust, és imprescindible que hi hagi un mecanisme d'assignació de la participació en el programa concret i/o que hi hagi suficient informació disponible sobre els *outcomes* i sobre característiques d'un conjunt de participants i no participants en el programa. Començar a dissenyar l'avaluació d'impacte en el moment en què es dissenya el programa permet implementar un procés d'assignació de la participació que sigui adequat per a l'avaluació d'impacte i permet planejar la recollida de dades necessàries per a la realització de l'avaluació, elements que incrementen considerablement tant la viabilitat com la robustesa de l'avaluació. Per altra banda, pensar en les decisions que es volen informar en el moment de dissenyar el programa augmenta la utilitat d'una futura avaluació d'impacte. Un exemple clar d'això es dona quan un mateix programa pot tenir diferents alternatives d'implementació i és rellevant conèixer quina d'elles és més efectiva per a la consecució d'un objectiu determinat.

EXEMPLES

► Avaluacions d'impacte utilitzant el mètode experimental:

Avaluació d'impacte del programa "Primer la llar" per a persones sense llar. Ivàlua, 2018.

https://ivalua.cat/sites/default/files/2022-04/Informe_Avaluacio_Impacte_PrimerLlar.pdf

Avaluació d'impacte del programa Èxit Estiu. Ivàlua, 2018.

https://ivalua.cat/sites/default/files/2020-06/Exit_informe_2018_o.pdf

► Avaluacions d'impacte utilitzant mètodes quasi-experimentals:

Avaluació d'impacte del programa Èxit Curs. Ivàlua, 2017.

https://ivalua.cat/sites/default/files/2020-01/Exit_Estiu_informe_gen_2017_def%20%281%29.pdf

Avaluació d'impacte dels Programes de Qualificació Professional Inicial. Ivàlua, 2013.

https://ivalua.cat/sites/default/files/2019-11/10_o3_2014_o8_41_24_Informe_PQPI_final_2013_o.pdf

PER SABER-NE MÉS

► Webs de recursos

UNICEF Impact Evaluation series:

<https://www.betterevaluation.org/tools-resources/unicef-impact-evaluation-series>

► Guies i manuals

Guia pràctica 5. Avaluació d'impacte. Ivàlua, 2009.

https://ivalua.cat/sites/default/files/2019-10/01_03_2010_11_33_12_Guia5_Impacte_Setembre2009_revfeb2010_massavermella.pdf

Guia pràctica 10. Avaluar l'impacte de les polítiques actives d'ocupació. Ivàlua, 2013.

https://ivalua.cat/sites/default/files/2019-10/17_12_2013_11_58_01_Guia10_PAO_cat.pdf

Guia pràctica 11. Com avaluar l'impacte de les polítiques educatives. Ivàlua, 2015.

https://ivalua.cat/sites/default/files/2019-10/08_06_2015_12_14_37_Guia11%20politica_educativa_CAT.pdf

Gertler, Paul J., Martinez, S., Premand, P., Rawlings, Laura B., Vermeersch, Christel M. J. (2016). La evaluación de impacto en la práctica, Segunda edición. Washington, DC: Inter-American Development Bank i World Bank.

<https://openknowledge.worldbank.org/server/api/core/bitstreams/6f2ee-bf7-1a3c-5f67-a9c3-c39f68299ed9/content>

Amb el suport de:



Generalitat de Catalunya
**Departament
de Drets Socials**